



**HAL**  
open science

## Two different combinations of RNA-binding domains determine the RNA-binding specificity of nucleolin.

H. Ginisty, F. Amalric, Philippe Bouvet

► **To cite this version:**

H. Ginisty, F. Amalric, Philippe Bouvet. Two different combinations of RNA-binding domains determine the RNA-binding specificity of nucleolin.. *Journal of Biological Chemistry*, 2001, 276, pp.14338-14343. hal-00023731

**HAL Id: hal-00023731**

**<https://hal.science/hal-00023731>**

Submitted on 31 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

## Two Different Combinations of RNA-binding Domains Determine the RNA Binding Specificity of Nucleolin\*

Received for publication, December 11, 2000, and in revised form, January 9, 2001  
Published, JBC Papers in Press, January 18, 2001, DOI 10.1074/jbc.M011120200

Hervé Ginisty, François Amalric, and Philippe Bouvet‡§

From the *Laboratoire de Pharmacologie et de Biologie Structurale, CNRS UMR 5089, 205 route de Narbonne, 31077 Toulouse Cedex, France and the ‡Ecole Normale Supérieure de Lyon, CNRS UMR 5665, 46 Allée d'Italie, 69364 Lyon Cedex 07, France*

**Nucleolin is an abundant nucleolar protein involved in several steps of ribosome biogenesis. The protein is highly conserved through evolution and possesses four RNA-binding domains (RBD), which are likely to determine its RNA binding specificity. Previous studies have shown that nucleolin interacts with two different RNA targets. The first is a small stem-loop structure, the nucleolin recognition element (NRE), found all along the pre-ribosomal RNA. The second is a short single-stranded RNA sequence, the evolutionary conserved motif (ECM), located five nucleotides downstream of the first processing site in the pre-ribosomal RNA 5' external transcribed spacer. Biochemical, genetic, and structural studies have shown that the first two RBD of nucleolin are necessary and sufficient for the specific interaction of nucleolin with the NRE motif. In this work, we have studied the interaction of nucleolin with the ECM sequence. Deletion and mutational analyses showed that all four RBDs of hamster nucleolin were required for the interaction with the ECM sequence. This RNA binding specificity is conserved between hamster and *Xenopus laevis*, whereas the *Xenopus* protein does not interact with the NRE. Nucleolin is the first example of a protein that requires four RBDs for its interaction with an RNA target, demonstrating that a single protein can use different combinations of RBD to interact specifically with several RNA sequences.**

Specific RNA-protein interactions play an important role in gene expression. One of the most common protein sequence motifs involved in these interactions is the RNA-binding domain (RBD),<sup>1</sup> also called the RNA recognition motif (RRM) (1, 2). A single or multiple RBD, often in combination with other domains involved in protein-protein interaction, are found in proteins that have diverse functions such as in pre-mRNA maturation and splicing (3, 4), hnRNA packaging (5), and mRNA stability and translation (6, 7).

\* This work is supported by grants from the Association pour la Recherche contre le Cancer (ARC), CNRS, and Fondation pour la Recherche Médicale (FRM). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 33-4-72-728016; Fax: 33-4-72-728080; E-mail: pbouvet@ens-lyon.fr.

<sup>1</sup> The abbreviations used are: RBD, RNA-binding domain; RNP, ribonucleoprotein; hnRNP, heteronuclear RNP; pre-rRNA, precursor ribosomal RNA; ETS, external transcribed spacer; ECM, evolutionary conserved motif; NRE, nucleolin recognition element; nt, nucleotide(s); PCR, polymerase chain reaction; PAGE, polyacrylamide gel electrophoresis; SELEX, systematic evolution of ligands by exponential enrichment.

The RBD domain is characterized by two highly conserved motifs called RNP-1 and RNP-2 found within a weakly conserved 80-amino acid sequence (1, 2). The crystal and solution structures of several RBDs have revealed a common structural organization composed of four antiparallel  $\beta$ -strands packed against two  $\alpha$ -helices (Refs. 8–10; 14). The RNP-1 octapeptide and RNP-2 hexapeptide are located in the  $\beta$ 1 and  $\beta$ 3 strands, respectively. Structures of RBD found in tandem show that the two independently folded domains do not interact with each other in Sex-lethal (11, 12), HuC (13), and nucleolin proteins (14). In contrast, in hnRNPA1, the two domains interact extensively (15, 16). Despite the highly conserved structure of the RBD, they show a remarkable ability to interact with diverse RNA sequences and structures. The tertiary structures of several RBD:RNA complexes are now available. These are the single RBDs from the U1A and U2B'' protein bound to a stem-loop structure (8, 17, 18), the two RBDs of Sex-lethal, hnRNPA1, and poly(A)-binding protein bound to a single-stranded RNA (12, 16, 19), and the first two RBD of nucleolin bound to a stem-loop structure (20). These different RNA-protein complexes revealed that the RBD uses the highly conserved RNP1 and RNP2 motifs in addition to the more highly divergent loops and linker regions specific to each RBD for the specific binding with their respective RNA target.

The presence of several RBD motifs (up to four domains in poly(A)-binding protein and nucleolin) might suggest that these proteins could interact with several RNA targets or that all RBD are required for the interaction with a single RNA or both. All three RBDs of U2AF<sup>65</sup> are required for high affinity binding to the polypyrimidine tract (4). The first two RBDs of poly(A)-binding protein and nucleolin bind specifically to polyadenylate RNA (21–23) and a short stem-loop structure (24, 25), respectively, as efficiently as full-length protein; the function of the last two RBD is not known.

Nucleolin is a major nucleolar protein involved in most steps of ribosome biogenesis (26–29). The interaction of nucleolin with pre-ribosomal RNA (30–32) is believed to play an important role in rRNA maturation and pre-ribosome assembly (33, 34). Studies of the specificity of the interaction of nucleolin with pre-rRNA have revealed that it interacts with two different RNA motifs. The nucleolin recognition element (NRE) motif, identified by SELEX, is a small stem-loop structure (31). Nucleolin binds tightly to similar stem-loop structures found all along the pre-rRNA (35). The second motif, the evolutionary conserved motif (ECM), was identified during studies on the function of nucleolin in the first processing step of ribosomal RNA (33, 34). The ECM is a short evolutionary conserved 11-nt sequence found 5 nt downstream of the cleavage site (36). The interaction of nucleolin with this sequence is absolutely required for the assembly of the processing complex and maturation of pre-rRNA at this site (33, 34).

Although the first two RBD of nucleolin are required for the specific interaction with the NRE motif (24, 25), it was of interest to determine which domains are involved in the binding of the ECM. In this work we show that all four RBD of nucleolin are required for interaction with the ECM. This interaction is conserved between hamster and *Xenopus* nucleolin. This is the first example of a protein that utilizes two different combinations of RBD to interact with two different RNA targets.

#### MATERIALS AND METHODS

**Plasmids Constructs and in Vitro RNA Transcription**—Mouse rDNA fragment from nt 645–1250 was amplified by PCR using the following oligonucleotides: 5'-ETS-645 (5'-ggaagatctgcgctgttctcactc-3') and 5'-ETS-1250 (5'-ggaattcaaaccttcaaccacccagcgcg-3'). The EcoRI and Bg/III sites used for the cloning in pSP72 (Promega) to give pSPETS<sub>645–1250</sub> are underlined. The BglIII site present in the resulting plasmid was removed. The final plasmid was called pSPETS<sub>645–1250ΔBglIII</sub>, and was linearized by enzyme BspEI (+677) for *in vitro* transcription with T7 RNA polymerase to give RNA<sub>645/677</sub>. The NRE cloned between the XbaI and HindIII sites of pSP64PA plasmid (Promega) encode a 68-nucleotide-long RNA. Labeled RNA was synthesized using [ $\alpha$ -<sup>32</sup>P]CTP in the transcription reaction. Unincorporated nucleotides were removed by gel filtration (G50, Amersham Pharmacia Biotech), and then the RNA was ethanol precipitated.

**Production of Recombinant Proteins**—The different nucleolin mutants were produced as described previously (25). Nucleolin mutants were generated by PCR using Vent<sup>TM</sup> DNA polymerase (New England Biolabs) and hamster nucleolin cDNA as template. PCR products contained an NdeI and a BamHI site at their 5'- and 3'-ends, respectively, for subcloning in the corresponding sites of pET15b plasmid (Novagen). The following oligonucleotides were used: R1N (5'-ccccatgaaatctgctcatggaac-3') and R2C (5'-ccggatccggtaccagtataagaaagaaac-3') for R12; R1N and R3C (5'-ccggatccggtaccgttgaactcaaacctgat-3') for R123; R2N (5'-ccccatgacacttttagcaaaaaa-3') and R4C (5'-ccggatccggtaccgcccagtcacaagtaac-3') for R234; R1N and R4C for R1234; and X.NdeM (5'-ggaattccatgctgcaaaacgcaaaaaa-3') and XR4C (5'-ccggatccggtaccggaatctctcttagg-3'). All proteins were quantified using Bradford reagent (Bio-Rad) and checked on SDS-PAGE. Each RNA-binding domain is defined here from the first amino acid of the  $\beta$ 1 sheet to the last amino acid of the  $\beta$ 4 sheet. For Chinese hamster nucleolin, these amino acid residues are: Asn<sup>308</sup>-Pro<sup>381</sup> for RBD1, Thr<sup>394</sup>-Gly<sup>465</sup> for RBD2, Thr<sup>486</sup>-Gly<sup>558</sup> for RBD3, Pro<sup>563</sup>-Pro<sup>646</sup> for RBD4.

The introduction of the RNP-1 mutations was achieved by PCR site-directed mutagenesis with the following oligonucleotides: R1(L/F, L/F)S and R1(L/F, L/F)NS for R1(LL)234; R2(L/I, L/F)S and R2(L/I, L/F)NS for R12(LL)34; R3(LL)S and R3(LL)NS for R123(LL)4NS; R4(LL)S and R4(LL)NS for R1234(LL); R1(LL)S and R1(LL)NS for X.L.R1(LL)234; X.L.R2(LL)S and X.L.R2(LL)NS for X.L.R1(LL)34; R3(L L)S and R3(LL)NS for X.L.R123(LL)4NS; R4(LL)S and R4(LL)NS for X.L.R1234(LL): (R1(L/F, L/F)S, 5'-ggaatt(a)tggttt(a)tgtgaccttgagtc-3'; R1(L/F, L/F)NS, 5'-ccac(t/a)aaace(t/a)aattctactttgacc-3'; R2(L/I, L/F)S, 5'-aaaggg(c/a)ttgctt(t/a)atgtaatttaagtc-3'; R2(L/I, L/F)NS, 5'-aat(t/a)aaagcaa(g/t)cccttactcttccc-3'; R3(LL)S, 5'-gggttagcgttaatagaatttgcct-3'; R3(LL)NS, 5'-tattaacgtaacccttagattgccc-3'; R4(LL)S, 5'-gggttagcttagtagactcaacagtg-3'; R4(LL)NS, 5'-tactaaacctaacctttagaggaacc-3'; X.L.R2(LL)S, 5'-aaagggctgcaactggtgagtttagcactgaag-3'; X.L.R2(LL)NS, 5'-ctcaaccagtcgaccctttattgatcatcat-3'). All recombinant plasmids were sequenced to confirm the presence of the mutations.

**Expression and Purification of Recombinant Proteins**—BL21(DE3)plysS was transformed with each recombinant pET15b plasmid. Cells grown at 37 °C in LB (100 mg/liter ampicillin, 20 mg/liter chloramphenicol) were induced with 1 mM isopropyl-1-thio-D-galactopyranoside for 4 h. Harvested cells were resuspended in buffer A (50 mM sodium phosphate, pH 8, 300 mM NaCl) and lysed by sonication. After centrifugation (30 min at 10,000 × g), the supernatant was recovered and gently mixed for 1 h at 4 °C after the addition of 1  $\mu$ l of Ni<sup>2+</sup>-nitrilotriacetic acid resin (Qiagen)/ml of initial culture. After three washes with buffer A and two with buffer B (50 mM sodium phosphate, pH 6, 300 mM NaCl, 10% glycerol), tagged peptide was eluted with buffer C (buffer B + 0.5 M imidazole). The supernatant was applied on a G25 column (NAP 5, Amersham Pharmacia Biotech) equilibrated with 100 mM KCl and 10 mM Tris-HCl, pH 7.5. Concentrations were estimated with Bradford reagent (Bio-Rad protein assay) and checked by SDS-polyacrylamide gel electrophoresis.

**RNA Binding Assay**—Gel retardation assay were performed by in-

cubating 10 fmol of labeled RNA in TMKC buffer (20 mM Tris, pH 7.4, 4 mM MgCl<sub>2</sub>, 200 mM KCl, 20% glycerol, 1 mM dithiothreitol, 0.5 mg/ml tRNA, 4  $\mu$ g/ml bovine serum albumin) with the indicated amount of protein for 30 min at room temperature. Then they were directly loaded on 8% polyacrylamide gel (acrylamide:bisacrylamide = 60:1) containing 5% glycerol in 0.5× TBE at room temperature. The gel was dried and submitted to an autoradiography. Quantification of RNA retardation was performed with a phosphorimager (Fuji X-BAS 1000).

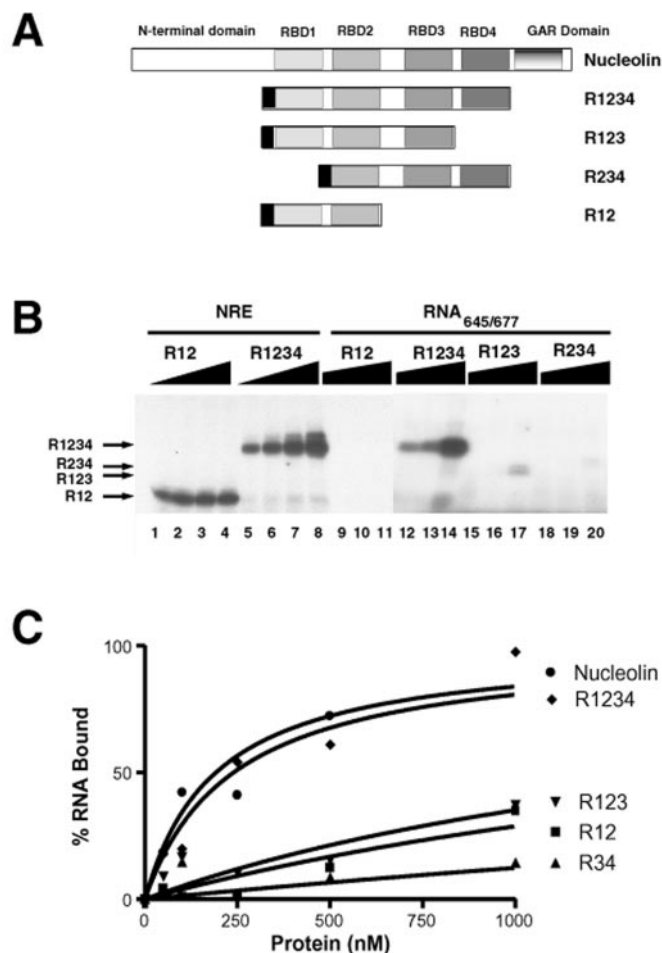
**UV Cross-link and Filter Binding Assays**—For the UV cross-linking assay, 500 fmol of labeled RNA were incubated with the indicated amount of protein at room temperature for 30 min in 10  $\mu$ l of TMKC buffer. Samples were subjected to UV light for 5 min as follows: 10- $\mu$ l drops were spotted on a thin layer of parafilm on ice and subjected to UV ( $\lambda$  = 254 nm) (Spectrolinker XL1000, Spectronics Corp.). Then 0.5  $\mu$ g of RNase A (Sigma) was added to each sample and incubated for 30 min at 37 °C. Proteins were resolved on SDS-PAGE.

Filter binding assays using purified nucleolin protein and *in vitro* labeled RNA were performed as described previously (25). Briefly, 10 fmol of labeled RNA were incubated with different concentrations of purified nucleolin or recombinant proteins in 50  $\mu$ l of binding buffer (200 mM KCl, 25 mM Tris, pH 7.5, 5 mM MgCl<sub>2</sub>, 20% glycerol, 75  $\mu$ g/ml poly(A), 10  $\mu$ g/ml bovine serum albumin) for 30 min at room temperature. Then, the reaction mixtures were filtered on a nitrocellulose membrane and washed three times with binding buffer. The percentage of bound RNA was determined using a phosphorimager.

#### RESULTS

**Nucleolin Utilizes Different RNA-binding Domains to Interact with Two Different RNAs**—We have previously shown that nucleolin interacts with two different RNA targets, the NRE and the ECM motifs (31, 34). The NRE motif is a short stem-loop structure found along the ribosomal RNA, whereas the ECM is an 11-nt motif located 5 nt downstream of the first processing site (34, 35). The first two RNA-binding domains (R12) of nucleolin are required for the interaction with the NRE sequence (24, 25). To determine whether the same RNA-binding domains were involved in the specific recognition of the ECM motif, we compared the interaction of the NRE and the ECM sequences with several recombinant proteins, which contain different numbers of nucleolin RBD (Fig. 1A). RNAs that contain the NRE or the ECM motif were *in vitro* transcribed in the presence of [ $\alpha$ -<sup>32</sup>P]CTP. The NRE RNA is 68 nt long and contains the stem-loop structure selected by SELEX (31). RNA<sub>645/677</sub> contains sequences 645–677 from mouse ribosomal RNA that include the conserved ECM motif. We have previously shown that interaction of nucleolin with this RNA requires the ECM sequence (34). These labeled RNAs were used in a UV cross-linking experiment in the presence of increasing amounts of R12 and R1234 proteins (Fig. 1B). After UV cross-linking and RNase digestion, proteins were resolved on SDS-PAGE. As previously described (24, 25), the NRE RNA interacts with R12 and R1234 with the same efficiency (Fig. 1B, lanes 1–8). In contrast, no cross-linking was observed between RNA<sub>645/677</sub> and R12 (lanes 9–11), whereas the R1234 protein interacts with this RNA (lanes 12–14). This experiment suggests that different combinations of RBDs are used for the specific interaction with the NRE and the ECM motifs.

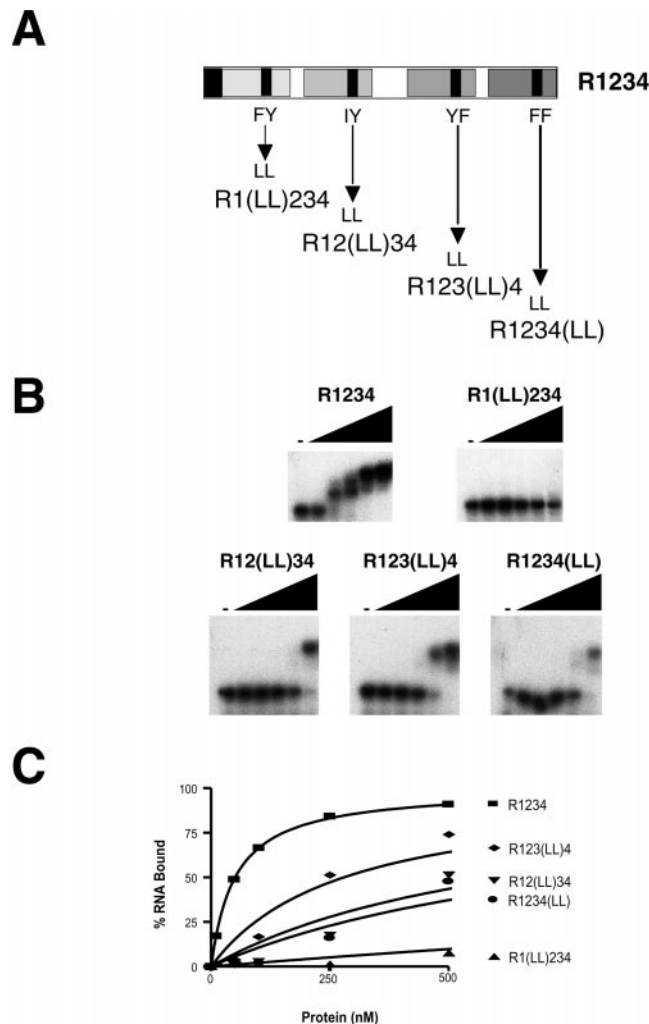
To determine which RBDs are involved in the interaction with RNA<sub>645/677</sub>, several recombinant proteins (see Fig. 1A) were produced in *Escherichia coli*, purified, and used in the UV cross-linking assay (Fig. 1B). Deletions of the first (R234, lanes 18–20) or the fourth (R123, lanes 15–17) RBD drastically reduced the cross-linking level compared with R1234 (lanes 12–14). To further confirm this experiment, we compared the interaction of full-length nucleolin with several recombinant proteins using a filter binding assay (Fig. 1C). Proteins and labeled RNAs were incubated at room temperature and filtered through a nitrocellulose filter. Only the RNA bound to the protein was retained on the filter, and the percentage of bound RNA could be determined using a phosphorimager. Nucleolin



**FIG. 1. Nucleolin utilizes different RBD combinations for the interaction with the NRE and ECM RNAs.** *A*, schematic representation of the nucleolin deletion mutants used in this study. RBDs are shown in gray boxes. The black boxes indicate the (his)<sub>6</sub> tag added at the N terminus of each peptide. The Gar domain is indicated by a striped box at the C-terminus of nucleolin. *B*, UV cross-linking assay. Different amounts of recombinant R12 (lanes 1–4 and 9–11), R1234 (lanes 5–8 and 12–14), R123 (lanes 15–17), or R234 (lanes 18–20) were incubated with labeled *in vitro* transcribed NRE (lanes 1–8) or RNA<sub>645/677</sub> (lanes 9–20). Protein concentrations are 125 nM (lanes 1 and 5), 250 nM (lanes 2, 6, 9, 12, 15, and 18), 500 nM (lanes 3, 7, 10, 13, 16, and 19), and 1000 nM (lanes 4, 8, 11, 14, 17, and 20). After RNase A digestion, labeled proteins were resolved on SDS-PAGE. *C*, filter binding assay. Purified nucleolin or recombinant proteins were incubated with labeled RNA<sub>645/677</sub> at room temperature. After filtration through a nitrocellulose filter, the retained, labeled RNAs (in complex with the protein) were quantified using a phosphorimager.

and the R1234 protein interact with the same affinity with RNA<sub>645/677</sub> ( $K_d$  of about 75 nM). Deletions of the last RBD (R123), the last two RBDs (R12), or the first two RBDs (R34) drastically reduced the interaction with this RNA. These experiments show that all RBD domains of nucleolin are required for the interaction with RNA<sub>645/677</sub>.

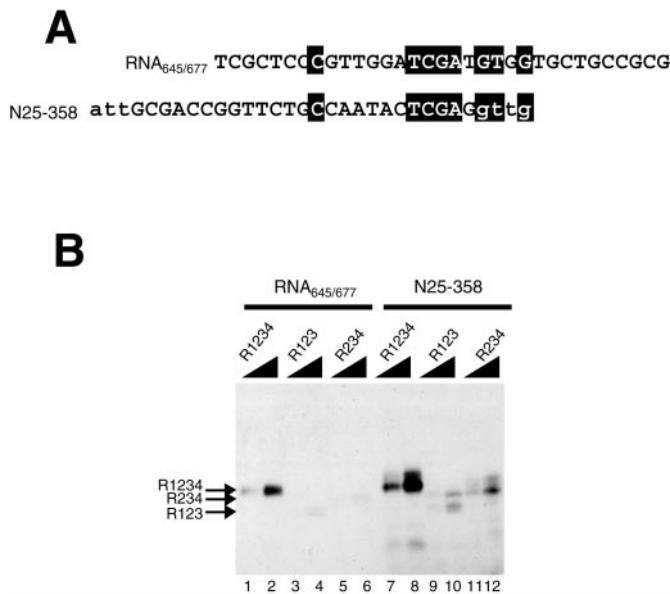
To further characterize the interaction of the four RBDs with this RNA target, point mutations were performed within each RBD. A characteristic feature of the RBD is the octamer RNP-1 motif that contains several aromatic residues involved in RNA binding. The presence of these aromatic residues is absolutely required for the binding activity of most RBDs (25, 37). The mutation of these amino acids to leucine residues in each RBD was performed (Fig. 2A), and each recombinant protein was purified and used in a gel shift assay (Fig. 2B). Quantification of these gel shifts using a phosphorimager is shown in Fig. 2C. Wild type protein (R1234) interacts with high binding affinity



**FIG. 2. Mutations in each RNP1 motif abolish interaction with the ECM RNA.** *A*, schematic representation of recombinant R1234 protein and of point mutations that were introduced in each RNP1 domain. The black box at the N terminus of the protein represents the (His)<sub>6</sub> tag used for the purification to homogeneity of each mutated protein. *B*, gel shift analysis with the wild type (R1234) and each mutated proteins with labeled RNA<sub>645/677</sub>. Protein concentrations were 0, 25, 50, 100, 250, 500, and 1000 nM for each protein. *C*, RNA binding curves for the different nucleolin mutants with RNA<sub>645/677</sub>. The legends for the data points are shown on the figure.

( $K_d$  50 nM) with RNA<sub>645/677</sub>. Mutation of the RNP1 motif in RBD1 and -4 (protein R1(LL)234 and R1234(LL), respectively), drastically reduced the interaction with the RNA ( $K_d > 500$  nM). This result is in agreement with data shown in Fig. 1, where deletions of either RBD1 or RBD2 completely abolished the interaction. Remarkably, mutations in RBD2 (R12(LL)34) or RBD3 (R123(LL)4) also drastically affect the interaction with the ECM. Altogether, these data demonstrate that all four RBDs are required for the specific interaction of nucleolin with RNA<sub>645/677</sub>.

**Interaction of Nucleolin with SELEX-selected Sequences**—A SELEX experiment performed with nucleolin (31) identified the NRE motif in 50% of the selected sequences. The remaining sequences show a significant homology to the ECM motif and are efficient competitors for the interaction of nucleolin with the ECM motif (34). To determine whether the interaction of nucleolin with the selected sequences that show homology to the ECM also required all four RBDs, one selected sequence (N25–358) was used in a UV cross-linking experiment (Fig. 3) with several recombinant proteins and was compared with the cross-linking results obtained with RNA<sub>645/677</sub>. Deletion of the



**FIG. 3. Interaction of hamster nucleolin with SELEX-selected sequences requires all four RBDs.** *A*, sequence alignment of RNA<sub>645/677</sub> with one of the SELEX-selected sequences (N25–358) that shows homology to the ECM motif (see Ref. 34 for the full alignment). The ECM motif is highlighted. The capital letters in N25–358 represent the 25 random sequences that have been selected by SELEX. *B*, UV cross-linking assay between RNA<sub>645/677</sub> or N25–358 and the R1234, R123, or R234 proteins. Increasing amounts of proteins, 250 nm in lanes 1, 3, 5, 7, 9, and 11 and 500 nm in lanes 2, 4, 6, 8, 10, and 12 were incubated with labeled RNA and subjected to UV cross-linking as described under “Experimental Procedures.” After RNase A digestion, proteins were resolved on SDS-PAGE. The arrows indicate the positions of the expected cross-link with each peptide.

first (Fig. 3*B*, R234, lanes 11 and 12) or the last (R123, lanes 9 and 10) RBD drastically reduced the interaction with RNA N25–358 compared with wild type R1234 protein (lanes 7 and 8). This experiment shows that the integrity of all four RBDs of nucleolin is required for interaction with the SELEX-selected sequence that presents homology with the ECM motif. It further suggests that the 11 nt found within this motif and partially in the selected sequence are key determinants for the specific interaction of the four RBD of nucleolin.

**Interaction of Nucleolin with the ECM Motif Is Conserved between Hamster and *Xenopus laevis***—Previous experiments have shown that human, mouse, and hamster nucleolin interacted with the same binding affinity and specificity with the NRE RNA (25, 35). This interaction requires the joint action of the first two RBD, which are well conserved in all these species. Although nucleolin RBDs are well conserved overall between human, mouse, and *X. laevis*, some significant differences can be found in RBD1 and -2 between human and *Xenopus* (only 53% identity to the hamster protein) (29). To determine whether the full-length or the first two RBD (*X.l.*R12) from *Xenopus* interact with the NRE motif, we performed a cross-link experiment with purified protein (data not shown) or recombinant *X.l.*R12 protein and compared it with the interaction of hamster R12 with the NRE (Fig. 4*A*). The full-length protein or the first two RBD of nucleolin did not interact with the NRE motif, indicating that key residues involved in the specific binding of the NRE are not conserved in the *Xenopus* protein. In contrast, nucleolin purified from *X. laevis* (34) or recombinant protein containing all four RBDs of the *Xenopus* protein (*X.l.*R1234) interacted with RNA<sub>645/677</sub> (Fig. 4*B*). This interaction required the ECM sequence, because deletion of the 11-nt motif completely inhibited the interaction with *X.l.*R1234 (Fig. 4*B*). To determine whether this interaction required all

four RBDs as for the hamster protein, mutations in the RNP-1 motif of each RBD of the *Xenopus* protein were performed. These recombinant proteins were purified and used in the cross-link assay (Fig. 4*C*). Mutations within the RNP-1 motif of RBD1 (R1(LL)234, lanes 4–6), RBD2 (R12(LL)34, lanes 7–9) and RBD4 (R1234(LL) lanes 13–15) had a drastic effect on the interaction with RNA<sub>645/677</sub>. The mutation in RBD3 (R123(LL)4) had a weaker effect (lanes 10–12), however, this is well correlated with the binding affinity of hamster R123(LL)4 protein (see Fig. 2*B*). The apparent higher interaction of *X.l.*R123(LL)4 with the ECM RNA might result from the cross-linking assay compared with gel shift analysis for the hamster protein.

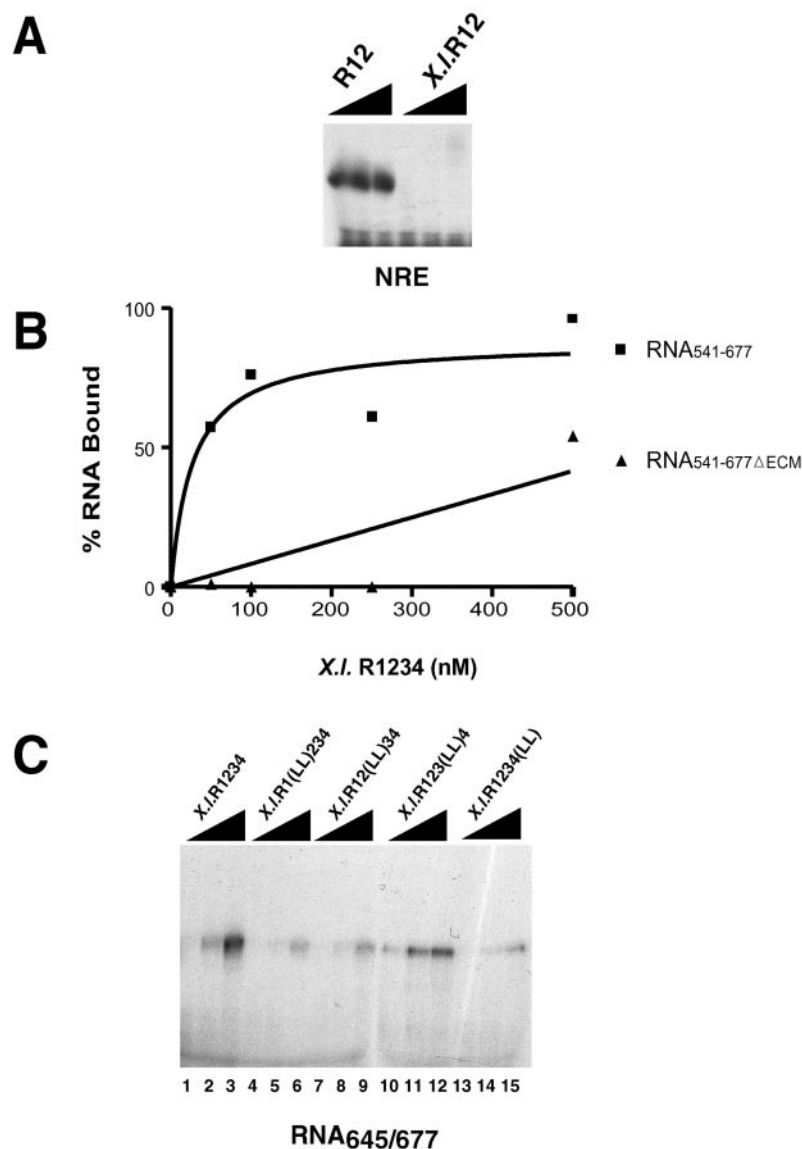
## DISCUSSION

In this work we show that all four RBDs of nucleolin are involved in the interaction with the evolutionary conserved RNA sequence (ECM) located 5 nt downstream of the first processing cleavage site. Deletion of the first or last RBD (Fig. 1) or mutation of residues within each RNP-1 motif (Fig. 2) drastically reduced interaction with the ECM. The ECM sequence is conserved between human, mice, and *X. laevis* (38, 39). Interestingly, nucleolin purified from *Xenopus* is able to activate the first processing step *in vitro* (34); the interaction of the *Xenopus* nucleolin with the ECM also requires all four RBD (Fig. 4). This conserved interaction of nucleolin with the ECM sequence through evolution indicates that nucleolin plays an important role in this first processing step.

**Nucleolin Interacts with Several RNA Targets**—A SELEX experiment performed with hamster nucleolin (31) has characterized several RNA binding sequences. Half of these sequences form a small stem-loop structure (NRE) with a short stem (5 base pairs) and a 7–10-nt loop containing the (U/G)CCCGA motif. Similar motifs are found all along pre-rRNA, and interaction of nucleolin with these sequences is believed to be involved in pre-ribosomal assembly (20, 35). Nucleolin of hamster, mouse, and human origin bind to the NRE motif with high affinity (25, 31, 35) ( $K_d$  5–20 nM). However, nucleolin from *X. laevis* does not bind with this RNA structure (Fig. 4). The first two RNA-binding domains of mouse nucleolin are necessary and sufficient for the high binding affinity and specific interaction with the NRE (24, 25). Interestingly, RBD1 and -2 of *Xenopus* nucleolin are only 53% identical to the hamster sequence, whereas human and mouse are about 85 and 96% identical, respectively, to the hamster protein (29). This finding suggests that key amino acids involved in the recognition of the NRE are not conserved within RBD1 and -2 of the *Xenopus* nucleolin. Indeed, structural analysis of the complex between hamster R12 and NRE has indicated that key amino acids residues involved in the interaction with the NRE RNA are not conserved in the *Xenopus* protein (20); this could explain why the *Xenopus* R12 protein does not interact with the NRE. Whether these RBDs are involved in the interaction with another target remains to be determined.

The remaining sequences selected by SELEX showed significant homology with the ECM motif found 5 nt downstream of the first processing site of rRNA (31, 34). We have shown that the specific interaction of nucleolin with the ECM motif is required for cleavage of the RNA 5 nt upstream of this motif *in vitro* (33, 34). The binding affinity of nucleolin with this RNA sequence ( $K_d$  100 nM) is lower compared with the interaction with the NRE sequence ( $K_d$  of 5–20 nM). Sequence analysis of the RNA sequence around the ECM predicts that this motif is in a single-stranded conformation (36, 39). Previous experiments suggest that the interaction of nucleolin with this sequence is required for the formation of the processing complex

**FIG. 4. RNA binding specificity of *Xenopus* nucleolin.** *A*, *Xenopus* nucleolin does not interact with the NRE motif. Hamster RBD12 (R12) and *Xenopus* nucleolin RBD12 (X.I.R12) were incubated with labeled NRE RNA, subjected to UV cross-linking, and then resolved on SDS-PAGE. *B*, filter binding assay with X.I.R1234 protein and labeled RNA<sub>541/677</sub> and RNA<sub>541/677</sub>ΔECM. *C*, effect of RNP1 mutations in each RBD of *Xenopus* nucleolin on the interaction with RNA<sub>645/677</sub>. Labeled RNA<sub>645/677</sub> was incubated with increasing amount of each recombinant protein (250, 500, or 1000 nM) and submitted to the UV-cross-linking assay as previously described.



by recruiting the different factors through interaction with the N- and C-terminal domains of nucleolin (33, 34).

**Multiple RBDs Are Involved in the Interaction with the ECM Sequence**—Deletion and mutational analyses of nucleolin RBDs (Figs. 2–4) show that all four nucleolin RBDs are involved in the interaction with the ECM sequence. Aromatic residues present in the RNP-1 sequence of the RBD contribute to RNA binding through ring-stacking interactions with the bases (9, 40). Mutation of these residues has been shown to drastically reduce the binding affinity of the RBD (37, 41). It is remarkable that mutations of these residues in each RBD of the hamster (Fig. 2) and *Xenopus* nucleolin (Fig. 4) drastically reduced the interaction with RNA<sub>645/677</sub>. This suggests that all four RBDs are involved in the interaction with the RNA and that this binding is conserved between hamster and *Xenopus*. Cooperation of several RBDs to achieve specific recognition of an RNA target seems to be a common feature of proteins that contains several RBDs. Two RBDs in hnRNP1, Sxl, ASF/SF2, HuD, HuC, and poly(A) binding protein are involved in binding with an RNA target (21, 40, 42–44). For U2AF<sup>65</sup>, three RBDs are required for high affinity binding to the polypyrimidine tract. (4) Nucleolin is the first example of a RNA-binding protein that requires four RBD for the specific interaction with an RNA sequence. The structural basis for the interaction of RBD

with a single-stranded RNA or a stem-loop structure has now been studied for a few RBD-RNA complexes. These structures are from U1A (8, 17, 45) and U2B'' (18), which bind a stem-loop structure, sxl (12); poly(A)-binding protein (19); and hnRNP1 (16), which binds single-stranded RNA and the structure of the first two RBD of nucleolin bound to the NRE motif (20). These different structures highlight the important role of the  $\beta$  sheet surface and loop regions within or outside the RBD to determine the specificity of the interaction. They also show that a common domain, which is highly conserved in numerous proteins, can adopt numerous conformations for specific recognition of many different RNA structures. Structural studies of the interaction of the four RBDs of nucleolin with the ECM RNA will certainly reveal a new mode of RNA protein interaction.

Another unique feature of nucleolin is its ability to use different combinations of RBD to interact with different RNAs. The interaction of nucleolin with the NRE and ECM motifs appears to be exclusive, because RBD1 and 2 are required for both interactions, and we have been unable to detect an interaction of both RNAs with the same nucleolin molecule (data not shown). The interaction of nucleolin with the two different RNA targets present on pre-rRNA suggests that nucleolin might play different roles upon interaction with pre-rRNA.

How nucleolin interaction with the NRE and ECM motifs is regulated remains to be determined, which will certainly reveal important aspects of protein dynamics. Recent studies of RNA-protein complexes show that interaction with RNA can induce conformational changes in the protein, which can influence protein function (45–47). The interaction of nucleolin with the different RNA targets might induce a selective conformational change of the protein allowing access of the N- and C-terminal domains, which are involved in protein-protein interactions. Depending on these interactions, nucleolin will be incorporated in different RNP complexes, which could explain its different functions.

## REFERENCES

- Birney, E., Kumar, S., and Krainer, A. R. (1993) *Nucleic Acids Res.* **21**, 5803–16
- Burd, C. G., and Dreyfuss, G. (1994) *Science* **265**, 615–621
- Sachs, A. B., Bond, M. W., and Kornberg, R. D. (1986) *Cell* **45**, 827–835
- Zamore, P. D., Patton, J. G., and Green, M. R. (1992) *Nature* **355**, 609–614
- Dreyfuss, G., Matunis, M. J., Pinol-Roma, S., and Burd, C. G. (1993) *Annu. Rev. Biochem.* **62**, 289–321
- Zhang, W., Wagner, B. J., Ehrenman, K., Schaefer, A. W., DeMaria, C. T., Crater, D., DeHaven, K., Long, L., and Brewer, G. (1993) *Mol. Cell. Biol.* **13**, 7652–7665
- Naranda, T., MacMillan, S. E., and Hershey, J. W. (1994) *J. Biol. Chem.* **269**, 32286–32292
- Allain, F. H., Gubser, C. C., Howe, P. W., Nagai, K., Neuhaus, D., and Varani, G. (1996) *Nature* **380**, 646–650
- Nagai, K., Oubridge, C., Jessen, T. H., Li, J., and Evans, P. R. (1990) *Nature* **348**, 515–520
- Gorlach, M., Burd, C. G., and Dreyfuss, G. (1994) *J. Biol. Chem.* **269**, 23074–23078
- Crowder, S. M., Kanaar, R., Rio, D. C., and Alber, T. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 4892–4897
- Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y., and Yokoyama, S. (1999) *Nature* **398**, 579–585
- Inoue, M., Muto, Y., Sakamoto, H., and Yokoyama, S. (2000) *Nucleic Acids Res.* **28**, 1743–1750
- Allain, F. H., Gilbert, D. E., Bouvet, P., and Feigon, J. (2000) *J. Mol. Biol.* **303**, 227–241
- Shamoo, Y., Krueger, U., Rice, L. M., Williams, K. R., and Steitz, T. A. (1997) *Nat. Struct. Biol.* **4**, 215–222
- Ding, J., Hayashi, M. K., Zhang, Y., Manche, L., Krainer, A. R., and Xu, R. M. (1999) *Genes Dev.* **13**, 1102–1115
- Oubridge, C., Ito, N., Evans, P. R., Teo, C. H., and Nagai, K. (1994) *Nature* **372**, 432–438
- Price, S. R., Evans, P. R., and Nagai, K. (1998) *Nature* **394**, 645–650
- Deo, R. C., Bonanno, J. B., Sonenberg, N., and Burley, S. K. (1999) *Cell* **98**, 835–845
- Allain, F. H., Bouvet, P., Dieckmann, T., and Feigon, J. (2000) *EMBO J.* **19**, 6870–6881
- Burd, C. G., Matunis, E. L., and Dreyfuss, G. (1991) *Mol. Cell. Biol.* **11**, 3419–3424
- Kuhn, U., and Pieler, T. (1996) *J. Mol. Biol.* **256**, 20–30
- Nietfeld, W., Mentzel, H., and Pieler, T. (1990) *EMBO J.* **9**, 3699–3705
- Bouvet, P., Jain, C., Belasco, J. G., Amalric, F., and Erard, M. (1997) *EMBO J.* **16**, 5235–5246
- Serin, G., Joseph, G., Ghisolfi, L., Bauzan, M., Erard, M., Amalric, F., and Bouvet, P. (1997) *J. Biol. Chem.* **272**, 13109–13116
- Busch, H. (1998) *J. Tumor Marker Oncol.* **12**, 5–68
- Olson, M. O. J. (1990) in *The Eukaryotic Nucleus—Molecular Biochemistry and Macromolecular Assemblies* (Strauss, P. R., and Wilson, S. M., eds) Vol. 2, pp. 519–559, Telford Press, Caldwell, NJ
- Srivastava, M., and Pollard, H. B. (1999) *FASEB. J.* **13**, 1911–1922
- Ginisty, H., Sicard, H., Roger, B., and Bouvet, P. (1999) *J. Cell Sci.* **112**, 761–772
- Ghisolfi, L., Kharrat, A., Joseph, G., Amalric, F., and Erard, M. (1992) *Eur. J. Biochem.* **209**, 541–548
- Ghisolfi-Nieto, L., Joseph, G., Puvion-Dutilleul, F., Amalric, F., and Bouvet, P. (1996) *J. Mol. Biol.* **260**, 34–53
- Herrera, A. H., and Olson, M. O. (1986) *Biochemistry* **25**, 6258–6264
- Ginisty, H., Amalric, F., and Bouvet, P. (1998) *EMBO J.* **17**, 1476–1486
- Ginisty, H., Serin, G., Ghisolfi-Nieto, L., Roger, B., Libante, V., Amalric, F., and Bouvet, P. (2000) *J. Biol. Chem.* **275**, 18845–18850
- Serin, G., Joseph, G., Faucher, C., Ghisolfi, L., Bouche, G., Amalric, F., and Bouvet, P. (1996) *Biochimie (Paris)* **78**, 530–538
- Craig, N., Kass, S., and Sollner-Webb, B. (1991) *Mol. Cell. Biol.* **11**, 458–467
- Merrill, B. M., Stone, K. L., Cobianchi, F., Wilson, S. H., and Williams, K. R. (1988) *J. Biol. Chem.* **263**, 3307–3313
- Mougey, E. B., Pape, L. K., and Sollner-Webb, B. (1993) *Mol. Cell. Biol.* **13**, 5990–5998
- Michot, B., and Bachelier, J. P. (1991) *Eur. J. Biochem.* **195**, 601–609
- Kanaar, R., Lee, A. L., Rudner, D. Z., Wemmer, D. E., and Rio, D. C. (1995) *EMBO J.* **14**, 4530–4539
- Mayeda, A., Munroe, S. H., Caceres, J. F., and Krainer, A. R. (1994) *EMBO J.* **13**, 5483–5495
- Burd, C. G., and Dreyfuss, G. (1994) *EMBO J.* **13**, 1197–1204
- Shamoo, Y., Abdul-Manan, N., Patten, A. M., Crawford, J. K., Pellegrini, M. C., and Williams, K. R. (1994) *Biochemistry* **33**, 8272–8281
- Tacke, R., and Manley, J. L. (1995) *EMBO J.* **14**, 3540–3551
- Varani, L., Gunderson, S. I., Mattaj, I. W., Kay, L. E., Neuhaus, D., and Varani, G. (2000) *Nat. Struct. Biol.* **7**, 329–335
- Mogridge, J., Legault, P., Li, J., Van Oene, M. D., Kay, L. E., and Greenblatt, J. (1998) *Mol. Cell* **1**, 265–275
- Zheng, N., and Gierasch, L. M. (1997) *Mol. Cell* **1**, 79–87

## Two Different Combinations of RNA-binding Domains Determine the RNA Binding Specificity of Nucleolin

Hervé Ginisty, François Amalric and Philippe Bouvet

*J. Biol. Chem.* 2001, 276:14338-14343.

doi: 10.1074/jbc.M011120200 originally published online January 18, 2001

---

Access the most updated version of this article at doi: [10.1074/jbc.M011120200](https://doi.org/10.1074/jbc.M011120200)

Alerts:

- [When this article is cited](#)
- [When a correction for this article is posted](#)

[Click here](#) to choose from all of JBC's e-mail alerts

This article cites 47 references, 17 of which can be accessed free at <http://www.jbc.org/content/276/17/14338.full.html#ref-list-1>